

WHAT IS CLAIMED IS:

1. A method for producing a nucleotide incorporating enzyme that incorporates a non-natural or rare nucleotide analogue, the method comprising:

(a) providing a plurality of nucleic acid segments, which nucleic acid

5 segments encode all or part of one or more parental nucleotide incorporating enzymes or a homologue thereof;

(b) identifying at least one non-natural or rare nucleotide analogue to be incorporated by the nucleotide incorporating enzyme, which non-natural or rare nucleotide analogue is incorporated by at least one of the one or more parental nucleotide incorporating enzymes at an efficiency of less than about 10% the efficiency of a naturally occurring nucleotide;

(c) diversifying the plurality of nucleic acid segments, thereby producing a library of nucleic acids encoding nucleotide incorporating enzyme variants; and

(d) identifying at least one nucleotide incorporating enzyme variant that incorporates the non-natural or rare nucleotide analogue at least about 10% as efficiently as a naturally occurring nucleotide.

2. The method of claim 1, wherein the non-natural or rare nucleotide analogue is incorporated less efficiently than inosine by a parental nucleotide incorporating enzyme.

20 3. The method of claim 1, wherein the non-natural or rare nucleotide analogue is incorporated less efficiently than 7-deaza dGTP by a parental nucleotide incorporating enzyme.

4. The method of claim 1, wherein the non-natural or rare nucleotide analogue is incorporated at an efficiency of less than about 5% the efficiency of a naturally occurring nucleotide by a parental nucleotide incorporating enzyme.

25 5. The method of claim 1, comprising identifying a nucleotide analogue selected from the group consisting of: a nucleotide derivatized with a functional group, a nucleotide derivatized with a methyl group, a nucleotide derivatized with a nitrile group,

a nucleotide comprising an unnatural base analogue, a nucleotide comprising a fluorescent label, a nucleotide comprising a ribose or deoxyribose analogue, and a nucleotide comprising an unnatural glycosidic linkage to a base.

6. The method of claim 5, wherein the nucleotide analogue is labeled at

5 the 2' or 3' hydroxyl position.

7. The method of claim 6, wherein the 2' or 3' hydroxyl group is derivatized by direct attachment of a functional group.

**8.** The method of claim 7, wherein the functional group is selected from a fluoro- group, an alkoxy- group, or an amine group.

9. The method of claim 5, wherein the nucleotide analogue is labeled at a position other than the 2' or 3' hydroxyl position.

**10.** The method of claim 5, wherein the nucleotide comprising a fluorescent label comprises a fluorescent group attached via a linker or a fluorescent base analogue.

15                   **11.** The method of claim 10, wherein the fluorescent group is attached by  
an amide linkage or an ester linkage.

**12.** The method of claim 5, wherein the nucleotide is derivatized by esterification or thioesterification of a phosphate.

13. The method of claim 5, wherein the nucleotide comprising a ribose  
20 or deoxyribose analogue comprises a cyclopentyl-derived sugar.

**14.** The method of claim 5, wherein the nucleotide comprising an unnatural glycosidic linkage to a base comprises a C-C linkage or a C-O linkage.

15. The method of claim 1, wherein the nucleic acid segments comprise  
RNA polynucleotides, DNA polynucleotides, or character strings representing RNA or  
25 DNA polynucleotides stored in a computer readable medium.

16. The method of claim 1, wherein the nucleic acid segments are produced by one or more of enzymatic digestion, chemical cleavage and mechanical fragmentation.

17. The method of claim 1, wherein the nucleic acid segments are 5 artificially synthesized.

18. The method of claim 17, wherein the artificially synthesized nucleic acid segments comprise oligonucleotides.

19. The method of claim 1, wherein the plurality nucleic acid segments encode a plurality of inactive nucleotide incorporating enzymes or homologues thereof.

10 20. The method of claim 1, wherein the plurality of nucleic acid segments comprise a plurality of codon altered nucleic acid segments.

21. The method of claim 1, wherein the plurality nucleic acid segments comprise at least one nucleic acid segment comprising one or more intron or intein.

15 22. The method of claim 1, wherein the plurality of nucleic acid segments comprise sequences of two or more members of a family of nucleotide incorporating enzymes.

23. The method of claim 1, wherein the nucleic acid segments comprise members of two or more families of nucleotide incorporating enzymes.

20 24. The method of claim 1, wherein the one or more parental nucleotide incorporating enzyme comprises a nucleic acid polymerase, a terminal transferase, a ligase or a telomerase.

25. The method of claim 24, wherein the one or more nucleic acid polymerase comprises a DNA polymerase or an RNA polymerase.

26. The method of claim 25, wherein the one or more nucleic acid 25 polymerase comprises a DNA polymerase or an RNA polymerase selected from the

group consisting of: a DNA dependent DNA polymerase, an RNA dependent DNA polymerase, a DNA dependent RNA polymerase, or an RNA dependent RNA polymerase.

**27.** The method of claim 25, wherein the one or more nucleic acid

5 polymerase comprises a thermostable DNA polymerase.

**28.** The method of claim 1, comprising diversifying the plurality of nucleic acid segments by at least one diversity generating procedure comprising recombining or mutating at least one nucleic acid segment.

**29.** The method of claim 28, comprising diversifying the plurality of

10 nucleic acid segments by recombining the plurality of nucleic acid segments in vitro, in vivo, or in silico.

**30.** The method of claim 28, comprising recursively recombining the plurality of nucleic acid segments in vitro, in vivo, or in silico.

**31.** The method of claim 28, comprising recombining the plurality of

15 nucleic acid segments by assembling synthetic oligonucleotides.

**32.** The method of claim 31, wherein the synthetic oligonucleotides are

joined using only a ligase.

**33.** The method of claim 28, comprising diversifying the at least one

nucleic acid segment by an error prone PCR.

20 **34.** The method of claim 33, further comprising recombining at least one amplified product produced by the error prone PCR.

**35.** The method of claim 28, comprising mutating the at least one nucleic

acid segment, thereby producing one or more mutated nucleic acid segments.

**36.** The method of claim 35, further comprising recombining the one or

25 more mutated nucleic acid segments.

37. The method of claim 1, wherein at least one of the nucleic acids encoding nucleotide incorporating enzyme variants further comprises a vector.

38. The method of claim 37, wherein the vector comprises a replicable vector.

5 39. The method of claim 1, comprising identifying the at least one nucleotide incorporating enzyme variant that incorporates the non-natural or rare nucleotide analogue by mass spectroscopy.

10 40. The method of claim 1, comprising identifying the at least one nucleotide incorporating enzyme variant that incorporates the non-natural or rare nucleotide analogue by one or more method selected from the group consisting of optical or fluorescent spectroscopy, radiometry, chromatography, gel electrophoresis, capillary electrophoresis, streptavidin binding, hybridization, fluorescent resonance energy transfer, fluorescent polarization, and pyrophosphate detection.

15 41. The method of claim 1, comprising identifying the at least one nucleotide incorporating enzyme by:

(i) transforming the library of nucleic acids into a population of host cells, growth of which host cells is dependent on the presence of at least one essential naturally occurring nucleotide in a first medium;

20 (ii) growing the population of transformed host cells in a second medium, which second medium lacks the at least one essential naturally occurring nucleotide of the first medium and comprises the non-natural or rare nucleotide analogue identified in step (b); and

25 (iii) identifying at least one surviving transformed host cell, thereby identifying a nucleic acid encoding a nucleotide incorporating enzyme variant, which nucleotide incorporating enzyme variant incorporates the non-natural or rare nucleotide analogue.

42. The method of claim 1, comprising identifying the at least one nucleotide incorporating enzyme variant in a high throughput assay format.

43. The method of claim 1, further comprising pre-screening the library of nucleic acids encoding nucleotide incorporating enzyme variants by:

- (i) transforming the library of nucleic acids into a population of temperature sensitive nucleotide incorporating enzyme deficient bacterial host cells;
- 5 (ii) growing the transformed bacterial host cells at the non-permissive temperature; and

- (iii) identifying one or more transformed bacterial host cells capable of growth at the non-permissive temperature, thereby identifying one or more members of the library of nucleic acids that encodes a functional nucleotide incorporating enzyme.

10 44. The method of claim 1, further comprising identifying at least one nucleotide incorporating enzyme variant with at least one additional desired property.

15 45. The method of claim 44, wherein the at least one additional desired property is selected from the group consisting of thermostability, evenness of nucleotide incorporation, efficient terminal transferase activity, low fidelity, high fidelity, processivity, strand-displacement activity, nick translation activity, exchange reaction, cation requirement, modulation of activity by cation, sulfhydryal reagent requirement, shelf life, salt tolerance, organic solvent tolerance, mechanical stress tolerance, tolerance to impurities, altered pH dependence, altered dependence on buffer conditions, template composition, primer composition, and improved stability.

20 46. The method of claim 44, comprising identifying at least one nucleotide incorporating enzyme variant by simultaneously screening for incorporation of the non-natural or rare nucleotide analogue and at least one other desired property.

47. A method for producing a nucleotide incorporating enzyme that incorporates a non-natural or rare nucleotide analogue, the method comprising:

25 (a) providing a plurality of nucleic acid segments, which nucleic acid segments encode all or part of one or more parental nucleotide incorporating enzymes or a homologue thereof;

(b) identifying at least one non-natural or rare nucleotide analogue to be incorporated by the nucleotide incorporating enzyme, which non-natural or rare

nucleotide analogue is incorporated by at least one of the one or more parental nucleotide incorporating enzymes at an efficiency of less than about 10% the efficiency of a naturally occurring nucleotide;

(c) diversifying the plurality of nucleic acid segments, thereby producing a

5 library of nucleic acids encoding nucleotide incorporating enzyme variants; and

(d) identifying at least one nucleotide incorporating enzyme variant that incorporates the non-natural or rare nucleotide analogue at least about 10 fold more efficiently than at least one of the one or more parental nucleotide incorporating enzyme.

48. The method of claim 47, wherein the non-natural or rare nucleotide

10 analogue is incorporated at an efficiency of less than about 5% the efficiency of a naturally occurring nucleotide.

49. The method of claim 47, wherein the non-natural or rare nucleotide

analogue is incorporated at an efficiency of less than about 1% the efficiency of a naturally occurring nucleotide.

50. The method of claim 47, wherein the non-natural or rare nucleotide

analogue is incorporated at an efficiency of less than about 0.5% the efficiency of a naturally occurring nucleotide.

51. The method of claim 47, wherein the non-natural or rare nucleotide

analogue is incorporated at an efficiency of less than about 0.1% the efficiency of a naturally occurring nucleotide.

52. The method of claim 47, wherein the non-natural or rare nucleotide

analogue is incorporated at an efficiency of less than about 0.05% the efficiency of a naturally occurring nucleotide.

53. The method of claim 47, wherein the non-natural or rare nucleotide

25 analogue is incorporated at an efficiency of less than about 0.01% the efficiency of a naturally occurring nucleotide.

54. The method of claim 47, wherein the at least one nucleotide incorporating enzyme variant identified in step (d) incorporates the non-natural or rare nucleotide analogue at least about 20 fold more efficiently than at least one of the one or more parental nucleotide incorporating enzyme.

5 55. The method of claim 47, wherein the at least one nucleotide incorporating enzyme variant identified in step (d) incorporates the non-natural or rare nucleotide analogue at least about 50 fold more efficiently than at least one of the one or more parental nucleotide incorporating enzyme.

10 56. The method of claim 47, wherein the at least one nucleotide incorporating enzyme variant identified in step (d) incorporates the non-natural or rare nucleotide analogue at least about 100 fold more efficiently than at least one of the one or more parental nucleotide incorporating enzyme.

15 57. The method of claim 1 or 47, further comprising extending a plurality of nucleic acid segments annealed to a single stranded template using the nucleotide incorporating enzyme variant identified in step (d).

58. The method of claim 1 or 47, further comprising performing at least one PCR using the nucleotide incorporating enzyme variant identified in step (d).

20 59. The method of claim 1 or 47, further comprising performing at least one sequencing reaction using the nucleotide incorporating enzyme variant identified in step (d).

60. A method for producing a nucleotide incorporating enzyme with increased tolerance to biological impurities, the method comprising:

25 (a) providing a plurality of nucleic acid segments, which nucleic acid segments encode all or part of one or more parental nucleotide incorporating enzymes or a homologue thereof;

(b) diversifying the plurality of nucleic acid segments, thereby producing a library of nucleic acids encoding nucleotide incorporating enzyme variants; and

PCT/US2014/042668  
10

(c) identifying at least one nucleotide incorporating enzyme variant that efficiently polymerizes a polynucleotide in a template dependent manner in the presence of a biological impurity found in blood, plasma or urine.

5       **61.** The method of claim 60, wherein the nucleotide incorporating enzyme is a DNA polymerase.

**62.** The method of claim 60, wherein the nucleotide incorporating enzyme is an RNA dependent DNA polymerase.

**63.** The method of claim 60, wherein the nucleotide incorporating enzyme is a reverse transcriptase.

10       **64.** The method of claim 60, comprising providing a plurality of nucleic acid segments comprising a parental polynucleotide incorporating enzyme from at least one thermophilic organism.

**65.** The method of claim 60, comprising providing a plurality of nucleic acid segments comprising at least one bacterial DNA Polymerase I.

15       **66.** The method of claim 60, wherein the at least one bacterial DNA Polymerase I is a *Thermus* DNA Polymerase I.

**67.** The method of claim 60, wherein the at least one nucleotide incorporating enzyme identified in step (c) is identified in a high throughput assay.

20       **68.** The method of claim 60, wherein the at least one nucleotide incorporating enzyme identified in step (c) comprises a thermostable enzyme.

**69.** The method of claim 60, wherein the at least one nucleotide incorporating enzyme identified in step (c) comprises an enzyme that is capable of incorporating dUTP.

25       **70.** The method of claim 60, wherein the at least one nucleotide incorporating enzyme identified in step (c) comprises an enzyme that is incorporates

dUTP at least as efficiently as a nucleotide incorporating enzyme selected from the group consisting of *T. aquaticus* DNA polymerase (Taq) and *T. thermophilus* DNA polymerase (Tth).

71. The method of claim 60, wherein the at least one nucleotide  
5 incorporating enzyme identified in step (c) is active in a reaction mixture comprising at least about 20% blood.

72. The method of claim 60, wherein the at least one nucleotide incorporating enzyme identified in step (c) is active in a reaction mixture comprising at least about 50% plasma.

10 73. The method of claim 60, wherein the at least one nucleotide incorporating enzyme identified in step (c) is active in a reaction mixture comprising at least about 50% urine.

74. A nucleotide incorporating enzyme variant produced by the method of claim 1, 47 or 60.

15 75. The nucleotide incorporating enzyme variant of claim 74, which nucleotide incorporating enzyme variant incorporates the non-natural or rare nucleotide analogue at least about 10% as efficiently as a naturally occurring nucleotide.

20 76. The nucleotide incorporating enzyme variant of claim 74, which nucleotide incorporating enzyme variant incorporates the non-natural or rare nucleotide analogue at least about 10 fold more efficiently than at least one of the one or more parental nucleotide incorporating enzymes.

77. The nucleotide incorporating enzyme variant of claim 74, which variant incorporates nucleotides or nucleotide analogues with low fidelity.

25 78. The nucleotide incorporating enzyme variant of claim 74, which variant incorporates nucleotides or nucleotide analogues with high fidelity.

79. A kit comprising the nucleotide incorporating enzyme variant of claim 74, and one or more of a container, a packaging material, and a natural nucleotide or non-natural or rare nucleotide analogue.

80. A kit comprising the nucleotide incorporating enzyme variant of  
5 claim 75, and one or more of a container, a packaging material, and a non-natural or rare nucleotide analogue, which non-natural or rare nucleotide analogue is incorporated by a parental nucleotide incorporating enzyme at an efficiency of less than 10% the efficiency of a naturally occurring nucleotide, and which non-natural or rare nucleotide analogue is incorporated by the nucleotide incorporating enzyme variant at least 10% as efficiently as a naturally occurring nucleotide.  
10

81. A kit comprising the nucleotide incorporating enzyme variant of claim 76, and one or more of a container, a packaging material, and a non-natural or rare nucleotide analogue, which non-natural or rare nucleotide analogue is incorporated by a parental nucleotide incorporating enzyme at an efficiency of less than 10% the efficiency of a naturally occurring nucleotide, and which non-natural or rare nucleotide analogue is incorporated by the nucleotide incorporating enzyme variant at least 10 fold more efficiently than it is incorporated by at least one of the one or more parental nucleotide incorporating enzymes.  
15

82. An integrated system comprising a non-natural nucleotide analogue,  
20 a nucleotide incorporating enzyme variant produced by the method of claim 1, 47, or 74 and a detector.

83. The integrated system of claim 82, further comprising one or more of a user input device, a data processing device, a data output device, and a robotic controller.

25 84. The method of claim 1, wherein the nucleic acid segments are oligonucleotides, and wherein the diversifying step (c) comprises:

- i) generating a plurality of partially duplexed oligonucleotides by hybridizing the oligonucleotides from step (a), said duplexed oligonucleotides having overhangs of unhybridized regions;
- ii) assembling the plurality of partially duplexed oligonucleotides by hybridizing the overhangs of two or more partially duplexed oligonucleotides together; and,
- iii) ligating the assembled oligonucleotides to produce a library of recombinant nucleic acids, optionally in the presence of a polymerase.

**85.** The method of claim 60, wherein the nucleic acid segments are oligonucleotides, and wherein the diversifying step (b) comprises:

- i) generating a plurality of partially duplexed oligonucleotides by hybridizing the oligonucleotides from step (a), said duplexed oligonucleotides having overhangs of unhybridized regions;
- ii) assembling the plurality of partially duplexed oligonucleotides by hybridizing the overhangs of two or more partially duplexed oligonucleotides together; and,
- iii) ligating the assembled oligonucleotides to produce a library of recombinant nucleic acids, optionally in the presence of a polymerase.

**86.** A method for identifying a nucleotide incorporating enzyme with having a desired property, the method comprising:

20 a) providing a plurality of partially duplexed oligonucleotides having overhangs of unhybridized regions, which oligonucleotides comprise a subsequence of a nucleic acid encoding a nucleotide incorporating enzyme;

b) assembling the plurality of partially duplexed oligonucleotides by hybridizing the overhangs of two or more partially duplexed oligonucleotides together;

c) ligating the assembled oligonucleotides to produce a library of recombinant nucleic acids, optionally in the presence of a polymerase;

25 d) expressing the recombinant nucleic acids to generate a library of nucleotide incorporating enzyme variants; and,

e) screening the library of nucleotide incorporating enzyme variants for one or more desired property.

87. The method of claim 86, wherein the one or more desired property is selected from the group consisting of: incorporation of rare or non-natural nucleotides, thermostability, evenness of nucleotide incorporation, efficient terminal transferase activity, low fidelity, high fidelity, processivity, strand-displacement activity, nick 5 translation activity, exchange reaction, cation requirement, modulation of activity by cation, sulphydral reagent requirement, shelf life, salt tolerance, organic solvent tolerance, mechanical stress tolerance, tolerance to impurities, altered pH dependence, altered dependence on buffer conditions, template composition, primer composition, and improved stability.